

AD-A163 783 A TEST FOR PROGRAM EFFECT IN THE ABSENCE OF A PROPER
CONTROL GROUP. (U) FLORIDA STATE UNIV TALLAHASSEE DEPT
OF STATISTICS E S TOLLAR SEP 85 FSU-STATISTICS-M710
UNCLASSIFIED ARD-19367.33-MA DAG29-82-K-0168 F/G 12/1

A TEST FOR PROGRAM EFFECT IN THE ABSENCE OF A PROPER
CONTROL GROUP. (U) FLORIDA STATE UNIV TALLAHASSEE DEPT
OF STATISTICS E S TOLLAR SEP 85 FSU-STATISTICS-M710
ARO-19367. 33-MA DAG29-82-K-0168 F/G 12/1

維

UNCLASSIFIED

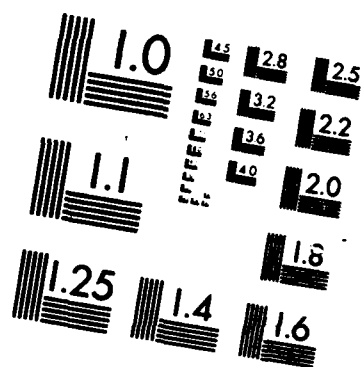
F/G 12/1

ML

END

PFI 5050

LTX₁



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963 A

2

SECURITY CLASSIFICATION

REPORT NUMBER

ARO 19367.33-MA

AD-A163 783

RECIPIENT'S CATALOG NUMBER

1. TITLE (and subtitle)

A TEST FOR PROGRAM EFFECT IN THE
ABSENCE OF A PROPER CONTROL GROUP

5. TYPE OF REPORT & PERIOD COVERED
Technical

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)

Eric S. Tollar

8. CONTRACT OR GRANT NUMBER(s)

USARO No. DAAG29-82-K-0168

9. PERFORMING ORGANIZATION NAME AND ADDRESS

The Florida State University
Department of Statistics
Tallahassee, FL 32306-3033

10. PROGRAM ELEMENT, PROJECT, TASK
AREA, & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME AND ADDRESS

U.S. Army Research Office - Durham
P.O. Box 12211
Triangle Park, NC 27709

12. REPORT DATE

September, 1985

13. NUMBER OF PAGES

16

14. PERFORMING AGENCY NAME & ADDRESS (if
different from Controlling Office)

15. SECURITY CLASS. (of this report)

15a. DECLASSIFICATION/DOWNGRADING
SCHEDULE

16. DISTRIBUTION STATEMENT (of this report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS

Analysis of Variance, Control Group, Convergence in distribution, t-distribution

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A statistical method is proposed to analyze certain designs for program effect when no proper control group is present. It is shown that under certain partially verifiable assumptions the data can be analyzed by standard analysis of variance techniques. If these assumptions are doubtful, then a statistic is proposed which is shown to have an asymptotic t distribution. The analysis is then carried out on some research data.

DTIC FILE COPY

DTIC
ELECTE
FEB 10 1986
S D

A TEST FOR PROGRAM EFFECT IN THE
ABSCENCE OF A PROPER CONTROL GROUP

by Eric S. Tollar

FSU Statistics Report No. M710
USARO Technical Report No. D-81

September, 1985

The Florida State University
Department of Statistics
Tallahassee, Florida 32306

Research supported by the U.S. Army Research Office under Grant No. DAAG-29-82-K-0168. The United States Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon.

Keywords: Analysis of Variance, Control Group, Convergence in distribution,
t distribution.

AMS (1980) Subject classifications: Primary 62F12 Secondary 62P99

A TEST FOR PROGRAM EFFECT IN THE
ABSCENCE OF A PROPER CONTROL GROUP

by

Eric S. Tollar

ABSTRACT

A statistical method is proposed to analyze certain designs for program effect when no proper control group is present. It is shown that under certain partially verifiable assumptions the data can be analyzed by standard analysis of variance techniques. If an assumption of no interaction is doubtful, then a new statistic is proposed which is shown to have an asymptotic t distribution. The analysis is then carried out on some research data.

Accession For	
NTIS CRA&I	<input checked="" type="checkbox"/>
ERIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution /	
Availability Codes	
Dist	Availability or Special
A-1	

1. INTRODUCTION

In the field of education, the experimenter often encounter a special problem in the statistical design of his experiment; the lack of a control group (for example, see Hersberger(1983)). Often, the scientist has devised a new educational program which he believes will augment or improve on the techniques presently used to teach skills of some sort to students. To test whether this new program is in fact an improvement, the scientist must approach an educational institute and obtain approval to teach his program to the target students. Since most of these programs appear beneficial to the students, the institute usually agrees to having the program implemented with one stipulation; that everyone who qualifies for the program is admitted into it. As such, the experimenter is not allowed to split the students into an experimental group and a control group, and the control group must be forgone. Often, nothing statistical can be salvaged from the data, but we suggest a statistic to measure the efficacy of the program for certain set-ups of the experiment when no control group is available.

The experimental set-up we address is one in which even in the experimental school, not all of the students available for the program actually qualify. This situation is especially common in the field of gifted or special education, where only those students who achieve some minimum (or maximum) score on an entrance exam are admitted. Further, we assume there are other schools available where the program will not be implemented, but the tests to determine admittance and to measure increases in learning skills can be administered. These schools we call the comparison schools.

Typically we see in the analysis of data by standard techniques that these other schools are pooled together and used as a control group. Often there is substantial evidence that these schools cannot be considered as members of the same population, and as such the pooling is improper. An allowance for the

"school" effect must be made. We deal with the case in which the "school" effect is clearly present, and the typical analysis is therefore inappropriate.

In section 2 we specify the statistical model used to describe the data, and show that frequently the data can be analyzed properly using standard analysis of variance techniques. In section 3 we deal with the case where the assumptions for the standard analysis of variance technique are not satisfied, and recommend an alternate statistic. It is shown that under reasonable assumptions this distribution will be approximately a t-distribution. In section 4 we analyze the data obtained by Hersberger (1983), using the proposed statistic.

2. THE MODEL

Let $\{X_{ijk}\}$ represent the final scores of the students (these scores we assume to be the only numbers of interest, they may be of the form post-test minus pre-test or other similar transformation of the data.) The value of i indicates the school of the student, so the range of i will be $0 \leq i \leq m$, where $i=0$ indicates a student from the program school. The value of j indicates if he qualified for the program ($j=1$), or did not qualify for the program ($j=0$), regardless of school. The value of k merely distinguishes which student in the various classes we have, where $1 \leq k \leq n_{ij}$.

We then assume for $0 \leq i \leq m$, $0 \leq j \leq 1$, $1 \leq k \leq n_{ij}$

$$X_{ijk} = \mu_i + \delta_i \cdot j + \epsilon_{ijk} \quad (2.1)$$

where $\epsilon_{ijk} \sim N(0, \sigma^2)$, i.i.d. for all i, j, k , and $\delta_i \sim F_i$, independent of $\{\epsilon_{ijk}\}$, and δ_j , $i \neq j$, for some distribution function F_i .

Except for the distribution of the δ_i 's, this is easily seen to be simply a reparametrization of the standard $2 \times m$ analysis of variance design. The δ_i 's measure the expected differences in the scores of the qualifying and non-qualifying students of the same school.

As a special case of the general model, we consider the case where $\delta_1 = \delta_2 = \dots = \delta_m = \delta$, for some constant δ . In this case, it is easy to see that $\delta_0 - \delta$ measures how much additional average difference qualifying students have over non-qualifying students when compared to qualifying students for the comparison schools. Unless there is a reason for believing the program school is special (without the presence of the program), the effect $\delta_0 - \delta$ must be attributed to the program. We therefore call $\delta_0 - \delta$ the program effect.

By straightforward analysis of variance techniques (see Graybill (1976)), it can be seen that the best linear unbiased estimate of $\delta_0 - \delta_1$ is

$$\hat{\delta}_0 - \hat{\delta}_1 = (\bar{x}_{01} - \bar{x}_{00}) - \left(\sum_{i=1}^m w_i \right)^{-1} \sum_{i=1}^m w_i (\bar{x}_{i1} - \bar{x}_{i0}), \quad (2.4)$$

where

$$\bar{x}_{ij} = n_{ij}^{-1} \sum_{k=1}^{n_{ij}} x_{ijk}, \quad (2.5)$$

and

$$w_i = n_{i0}n_{i1}(n_{i0} + n_{i1})^{-1}, \quad 0 \leq i \leq m. \quad (2.6)$$

Further, for

$$\hat{\sigma}^2 = \left[\sum_{i=0}^m (n_{i0} + n_{i1}) - 2(m+1) \right]^{-1} \sum_{i=0}^m [(n_{i0} - 1) s_{i0}^2 + (n_{i1} - 1) s_{i1}^2],$$

where

$$s_{ij}^2 = (n_{ij} - 1)^{-1} \sum_{k=1}^{n_{ij}} (x_{ijk} - \bar{x}_{ij})^2,$$

then it can be easily established that

$$(\hat{\delta}_0 - \hat{\delta}_1) \hat{\sigma}^{-1} [w_0^{-1} + (\sum_{i=1}^m w_i)^{-1}]^{-1/2} \quad (2.7)$$

has a t-distribution with noncentrality parameter

$$\lambda = (\delta_0 - \delta) \sigma^{-1} [w_0^{-1} + (\sum_{i=1}^m w_i)^{-1}]^{-1/2},$$

and degrees of freedom

$$v = (\sum_{i=0}^m n_{i0} + n_{i1}) - 2(m+1).$$

Of course, this test statistic can be implemented by most statistical computer packages by simply specifying the contrast $\delta_0 - m^{-1} \sum_{i=1}^m \delta_i$.

To test the assumption that $\delta_1 = \delta_2 = \dots = \delta_m = \delta$, one could temporarily delete the program school from the data set and run an analysis of variance with a high significance (say $\alpha = .5$ or $.75$) to test for interaction. If there is no evidence of interaction, it seems reasonable to accept that $\delta_1 = \delta_2 = \dots = \delta_m$. The statistic in (2.7) could then be used to test for significance of the program and give an estimate of the quantitative average effect of the program.

For the data of Hersberger(1983), it was found (perhaps surprisingly) that on the majority of measures there was no interaction in the control schools, and as such there is reason to believe the simple test procedure outlined above is appropriate in numerous experiments of the prescribed form.

3. THE STATISTIC WHEN INTERACTION IS PRESENT

Of course, occasionally there will be evidence of interaction at an unacceptable significance level. In such a case, the simple procedure outlined above is not applicable, and we must consider the more general model. Even in this case, we consider only a special case of the model (2.1).

We assume that $\delta_1, \dots, \delta_m$ is a random sample from some normally distributed population. As such, we have

$$\delta_i \sim N(\delta, n^2), \quad 1 \leq i \leq m. \quad (3.1)$$

We assume that δ_0 is from a normal population with the same variance and arbitrary mean, that is

$$\delta_0 \sim N(\gamma, n^2). \quad (3.2)$$

Of interest is $\gamma - \delta$; that is, are δ_0 and $\{\delta_1, \dots, \delta_m\}$ all elements of the same population? We will call $\gamma - \delta$ the expected program effect, since it is the additional difference expected in the program school scores.

In this case, we propose a new statistic, given by

$$t^* = (y_0 - \delta')[(m-1)^{-1} \sum_{i=1}^m w_i^* (y_i - \delta')^2]^{-1/2} [w_0^{*-1} + (\sum_{i=1}^m w_i^*)^{-1}]^{-1/2} \quad (3.3)$$

where

$$y_i = \bar{x}_{i1} - \bar{x}_{i0}, \quad (3.4)$$

$$w_i^* = (\hat{n}^2/\hat{\sigma}^2 + n_{i0}^{-1} + n_{i1}^{-1})^{-1}, \quad (3.5)$$

$$\hat{\sigma}^2 = (\sum_{i=0}^m \sum_{j=0}^1 (n_{ij} - 1) s_{ij}^2) / \sum_{i=0}^m \sum_{j=0}^1 (n_{ij} - 1), \quad (3.6)$$

$$\hat{n}^2 = [(m-1)^{-1} \sum_{i=1}^m (y_i - \bar{y})^2 - \hat{\sigma}^2 m^{-1} \sum_{i=1}^m (n_{i0}^{-1} + n_{i1}^{-1})]^{-1}, \quad (3.7)$$

and

$$\delta' = \sum_{i=1}^m w_i^* y_i / \sum_{i=1}^m w_i^*. \quad (3.8)$$

The main theorem of this paper, dealing with the asymptotic distribution of t^* , we now state.

THEOREM 3.1. Under the assumptions (2.1), (3.1), (3.2), as

$$\min_{\substack{0 \leq i \leq m \\ 0 \leq j \leq 1}} (n_{ij}) \rightarrow \infty, P(t^* \leq x) \rightarrow P(t \leq x),$$

where t has a t-distribution with $m-1$ degrees of freedom and non-centrality parameter $\lambda = (\gamma - \delta)/\eta\sqrt{1+m^{-1}}$.

Before proving the theorem we establish some intermediate results, which in addition to the proof of the theorem will shed light on the motivation for the statistic t^* . The first result is stated without proof as it is an elementary result which can be easily established using standard techniques (see Graybill (1976)).

LEMMA 3.2. If Z_1, \dots, Z_m are independent, normally distributed random variables with mean δ and variance σ^2/w_i respectively, then

- (1) the BLUE for δ is $\hat{\delta} = (\sum w_i)^{-1} (\sum w_i z_i)$, and
- (2) $\sigma^{-2} \sum w_i (z_i - \hat{\delta})^2 = \sigma^{-2} [\sum w_i^2 z_i^2 - (\sum w_i)^{-1} (\sum w_i z_i)^2]$ has a $\chi^2_{(m-1)}$ distribution, independent of $\hat{\delta}$.

Thus, to understand the motivation of the statistic t^* , consider the quantities $y_i = \bar{x}_{i1} - \bar{x}_{i0}$. It is clear that

$$y_i \sim N(\delta, \eta^2 + \sigma^2(n_{i1}^{-1} + n_{i0}^{-1})), \quad 1 \leq i \leq m, \quad \text{and} \quad (3.9)$$

$$y_0 \sim N(\gamma, \eta^2 + \sigma^2(n_{i1}^{-1} + n_{i0}^{-1})). \quad (3.10)$$

If $\eta^2 \sigma^{-2}$ were known, then letting

$$w_i^{-1} = \eta^2 \sigma^{-2} + (n_{i1}^{-1} + n_{i0}^{-1}), \quad (3.11)$$

from the previous lemma the statistic

$$(3.12) \quad (y_0 - \hat{\delta}) / [(\sum_{i=1}^m w_i (y_i - \hat{\delta})^2 / (m-1)) (w_0^{-1} + (\sum_{i=1}^m w_i^{-1})^{-1})]^{1/2}$$

would then have a t-distribution with $m-1$ degrees of freedom. The idea of t^* is to estimate η^2/σ^2 and use these to provide an estimate of the w_i 's. We know that σ^2 has a very reliable estimator available, $\hat{\sigma}^2$ given in equation (3.6).

It can be quickly determined that for $\bar{y} = m^{-1} \sum_{i=1}^m y_i$,

$$E(\sum_{i=1}^m (y_i - \bar{y})^2) = (m-1)\eta^2 + \frac{m-1}{m} \sigma^2 \sum_{i=1}^m (n_{i1}^{-1} + n_{i0}^{-1}). \quad (3.1)$$

Therefore, we can use as an estimate of η^2 the statistic $\hat{\eta}^2$ given in equation (3.7), and using w_i^* in equation (3.5) as an estimate of w_i in equation (3.12) we find the resultant statistic is t^* .

To establish that t^* is indeed asymptotically t-distributed, we need the following lemma.

LEMMA 3.3. If $\{X_n: n=1, 2, \dots\}$ is a sequence of random variables where

- 1) for some measure μ , X_n has density f_n with respect to μ
- 2) $f_n \leq k \cdot g$ for all n , and some density g , and
- 3) $f_n \rightarrow f$, for f a density,

then for any function $h(x,y)$ where $h(x,y)$ is continuous with respect to y for μ -almost all x , and for any function $\delta(x)$ where $\delta(X_n) \xrightarrow{P} \delta_0$,

$h(X_n, \delta(X_n)) \xrightarrow{d} h(X, \delta_0)$, where X has density f .

Remark: Condition 3) can be weakened to $X_n \xrightarrow{d} X$, some random variable, and the result is still true. It should be pointed out that conditions 1) and 2) coupled with $X_n \xrightarrow{d} X$ does not imply $f_n \rightarrow f$. However, this weakening only complicates the proof of the lemma, and since theorem 3.1 only requires the lemma as stated, we

do not prove the lemma under the weaker condition.

PROOF OF LEMMA 3.3. By the dominated convergence theorem (see Chung (1974)) we have that

$$\lim_{n \rightarrow \infty} \int e^{i h(x, \delta_0)} f_n(x) d\mu(x) = \int e^{i h(x, \delta_0)} f(x) d\mu(x),$$

and as such, $h(X_n, \delta_0) \xrightarrow{d} h(X, \delta_0)$. Therefore, to complete the proof, we need only show that $h(X_n, \delta(X_n)) - h(X_n, \delta_0) \xrightarrow{P} 0$. For $\gamma > 0$, let

$$A(\gamma) = \{x: \|h(x, \delta) - h(x, \delta_0)\| > \epsilon, \text{ some } \delta \text{ where } \|\delta - \delta_0\| \leq \gamma\}. \quad (3.14)$$

Since $h(x, \delta)$ is continuous with respect to δ , as $\gamma \rightarrow 0$, $A(\gamma) \downarrow A$, where $\mu(A) = 0$.

Let Y be a random variable with density g of condition 2). Then

$\lim_{\gamma \rightarrow 0} P(Y \in A(\gamma)) = 0$. As such, for all $\epsilon > 0$, there is a $\gamma_0 > 0$ where for all

$\gamma < \gamma_0$, $P(Y \in A(\gamma)) < \epsilon/k$. Thus, for all n , if $\gamma < \gamma_0$ then

$$P(X_n \in A(\gamma)) \leq k P(Y \in A(\gamma)) < \epsilon. \quad (3.15)$$

Choose an N where for $n > N$, $\gamma < \gamma_0$, $P(\|\delta(X_n) - \delta_0\| > \gamma) < \epsilon$. Then we have that

$$P(\|h(X_n, \delta(X_n)) - h(X_n, \delta_0)\| > \epsilon) \leq P(X_n \in A(\|\delta(X_n) - \delta_0\|)). \quad (3.16)$$

Observing that

$$P(X_n \in A(\|\delta(X_n) - \delta_0\|)) \leq P(X_n \in A(\|\delta(X_n) - \delta_0\|), \|\delta(X_n) - \delta_0\| \leq \gamma) + \epsilon \quad (3.17)$$

$$\leq P(X_n \in A(\gamma)) + \epsilon < 2\epsilon,$$

we have that $h(X_n, \delta(X_n)) - h(X_n, \delta_0) \xrightarrow{P} 0$, which completes the proof of the lemma. \square

We now proceed to establish theorem 3.1 as a consequence of lemma 3.3.

PROOF OF THEOREM 3.1. Let us first establish that for w_i^* as in (3.5), we have

$$\left(\sum_{i=1}^m w_i^* \right)^{-1} (w_0^*, w_1^*, \dots, w_m^*) \xrightarrow{P} \left(\frac{1}{m}, \frac{1}{m}, \dots, \frac{1}{m} \right). \quad (3.18)$$

First observe that for independent y_i where $y_i \sim N(\gamma, \eta^2 + \sigma^2(\frac{1}{n_{i1}} + \frac{1}{n_{i0}}))$, and for the density of $g(x)$ of $y \sim N(\gamma, \eta^2 + 2\sigma^2)$, $g(x)\sqrt{1+2\sigma^2/\eta^2}$ clearly dominates the density of each y_i . Also, as $\min_{\substack{0 \leq i \leq m \\ 0 \leq j \leq 1}} (n_{ij}) \rightarrow \infty$, each density clearly converges to

that of $N(\gamma, \eta^2)$. Thus, by the previous lemma (or more simply by the Skorohod convergence theorem, see Serfling (1980), we have that $\sum_{i=1}^m (y_i - \bar{y})^2 \xrightarrow{d} \eta^2 \chi^2(m-1)$.

Also, since we have that $\hat{\sigma}^2 \xrightarrow{P} \sigma^2$, we get from Slutsky's theorem (see Chung (1974)) that $\hat{\eta}^2/\hat{\sigma}^2 \xrightarrow{d} \frac{\eta^2}{\sigma^2(m-1)} \chi^2(m-1)$. We can therefore find an N_1 where

$P(\hat{\eta}^2/\hat{\sigma}^2 \leq \delta) < \epsilon$ for $\min_{\substack{0 \leq i \leq m \\ 0 \leq j \leq 1}} (n_{ij}) > N_1$.

Letting $c_i = (n_{i0}^{-1} + n_{i1}^{-1})^{-1}$, we can establish that

$$\frac{[\max_{1 \leq i \leq m} (c_i)]^{-1} - c_i^{-1}}{m(\hat{\eta}^2/\hat{\sigma}^2) + c_i^{-1}} \leq \frac{w_i^*}{\sum_{i=1}^m w_i^*} - \frac{1}{m} \leq \frac{[\min_{1 \leq i \leq m} (c_i)]^{-1} - c_i^{-1}}{m((\hat{\eta}^2/\hat{\sigma}^2) + c_i^{-1})}, \quad (3.18)$$

for $0 \leq i \leq m$.

Therefore

$$\left| \frac{w_i^*}{\sum_{i=1}^m w_i^*} - \frac{1}{m} \right| \leq \frac{[\min_{1 \leq i \leq m} (c_i)]^{-1} - [\max_{1 \leq i \leq m} (c_i)]^{-1}}{m \hat{\eta}^2/\hat{\sigma}^2}, \quad (3.19)$$

so we have that for $\min_{\substack{0 \leq i \leq m \\ 0 \leq j \leq 1}} (n_{ij}) > N_1$,

$$P\left(\left| \frac{w_i^*}{\sum_{i=1}^m w_i^*} - \frac{1}{m} \right| \leq \frac{\min_{1 \leq i \leq m} (c_i)^{-1} - \max_{1 \leq i \leq m} (c_i)^{-1}}{m\delta} \right) > 1 - \epsilon. \quad (3.20)$$

From (3.20), it is clear that $\left(\frac{w_0^*}{\sum_{i=1}^m w_i^*}, \frac{w_1^*}{\sum_{i=1}^m w_i^*}, \dots, \frac{w_m^*}{\sum_{i=1}^m w_i^*} \right) \xrightarrow{P} \left(\frac{1}{m}, \frac{1}{m}, \dots, \frac{1}{m} \right)$.

For $\underline{x} = (x_0, \dots, x_m)$, $\underline{\delta} = (\delta_0, \dots, \delta_m)$, $\hat{x} = \sum_{i=1}^m \delta_i x_i$, the function $h(\underline{x}, \underline{\delta}) =$

$$h(\underline{x}, \underline{\delta}) = \frac{x_0 - \hat{x}}{\left[\sum_{i=1}^m \delta_i (x_i - \hat{x})^2 / (m-1) \right]^{1/2} \left[\frac{1}{\delta_0} + 1 \right]^{1/2}}$$

is clearly continuous for all fixed \underline{x} with respect to $\underline{\delta}$. Letting $\underline{Y}_n = (y_0, \dots, y_m)$,

$\delta(\underline{Y}_n) = \left(\frac{w_0^*}{\sum_{i=1}^m w_i^*}, \dots, \frac{w_m^*}{\sum_{i=1}^m w_i^*} \right)$, we have that $h(\underline{Y}_n, \delta(\underline{Y}_n)) \xrightarrow{d} h(\underline{X}, (\frac{1}{m}, \frac{1}{m}, \dots, \frac{1}{m}))$,

where $X_i \sim N(\delta, \sigma^2)$, $i \geq 1$, $X_0 \sim N(\gamma, \sigma^2)$, and X_i 's are independent.

Clearly $h(\underline{X}, (\frac{1}{m}, \dots, \frac{1}{m})) \sim t(m-1)$, which complete the proof. \square

In the following section, we examine the data of Hersberger (1983), to show how the techniques in this and the preceding section can be used to analyze the effects of a program administered under the prescribed set-up.

4. ANALYSIS OF DATA

In 1983, Hersberger administered a program to measure the effectiveness of computers in the teaching of mathematics to students of high and low mathematical

ability (as defined by those students who scored in the upper 1/3 and lower 2/3 on a math pre-test, respectively). While we certainly do not analyze all the data of the study, we do select some particular measurements to illustrate the application of the analysis outlined above. The two measures selected are a computational measure, and a value of mathematics measure.

Below is a listing of the class means, standard deviations, and sample sizes for the computational measure, with a corresponding plotting of the class means.

Table 1. Summary of scores on the computational measure

Mathematical Ability	School			
	Program School	Comparison School 1	Comparison School 2	Comparison School 3
High	$\bar{x} = 34.67$	$\bar{x} = 29.96$	$\bar{x} = 24.00$	$\bar{x} = 28.43$
	$s = 5.19$	$s = 4.36$	$s = 4.81$	$s = 2.76$
	$n = 15$	$n = 9$	$n = 10$	$n = 7$
Low	$\bar{x} = 18.97$	$\bar{x} = 24.10$	$\bar{x} = 17.76$	$\bar{x} = 22.31$
	$s = 5.66$	$s = 5.55$	$s = 5.18$	$s = 6.05$
	$n = 37$	$n = 20$	$n = 17$	$n = 13$

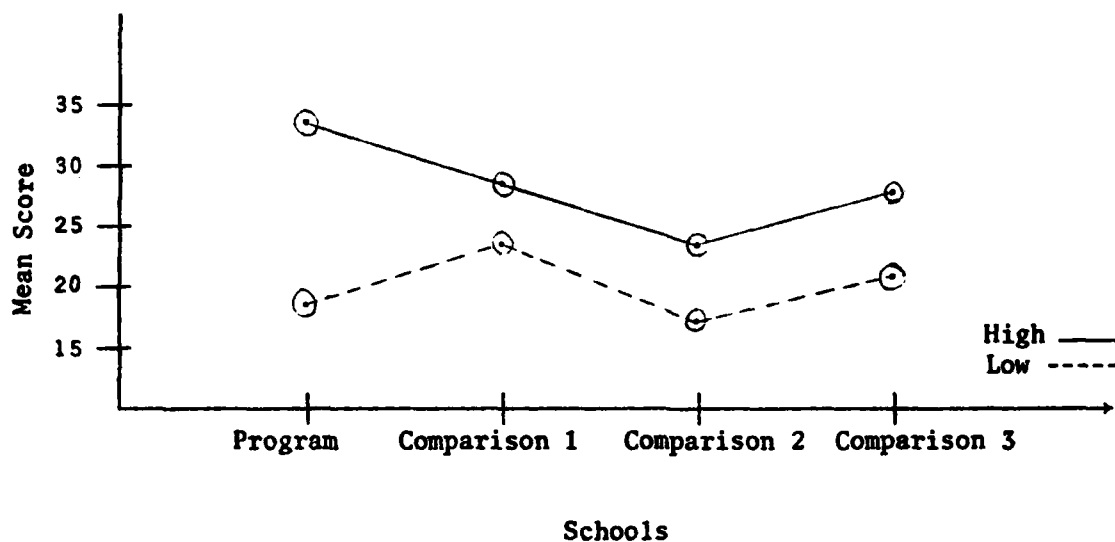


Figure 1. Cell means for the computational measure.

A test for interaction with the program school deleted yields an F-value of .04, which has a level of significance of .96. As such, the procedure outlined in section 2 using a conventional analysis is appropriate. From equations (2.4) and (2.7) we find that

$$\hat{\delta}_0 - \hat{\delta}_1 = 9.7759 ,$$

$$\hat{\sigma}^2 = 28.2516 , \text{ and}$$

$$t = 4.7123 \text{ with } df=120.$$

Thus, we find that the program effect is significant at the $p = .000006$ level, and conclude the program has a significant effect on computational skills.

The next measure we consider is the value of mathematics measure. Below is a listing of the class means, standard deviations and sample sizes, with a plotting of the class means.

Table 2. Summary of scores for value of mathematics measure

Mathematical Ability	School			
	Program School	Comparison School 1	Comparison School 2	Comparison School 3
High	$\bar{x} = 42.69$	$\bar{x} = 39.20$	$\bar{x} = 40.30$	$\bar{x} = 44.86$
	$s = 5.79$	$s = 5.31$	$s = 3.27$	$s = 2.80$
	$n = 16$	$n = 10$	$n = 10$	$n = 7$
Low	$\bar{x} = 39.28$	$\bar{x} = 40.00$	$\bar{x} = 37.00$	$\bar{x} = 41.92$
	$s = 5.73$	$s = 3.47$	$s = 6.60$	$s = 4.03$
	$n = 39$	$n = 23$	$n = 16$	$n = 13$

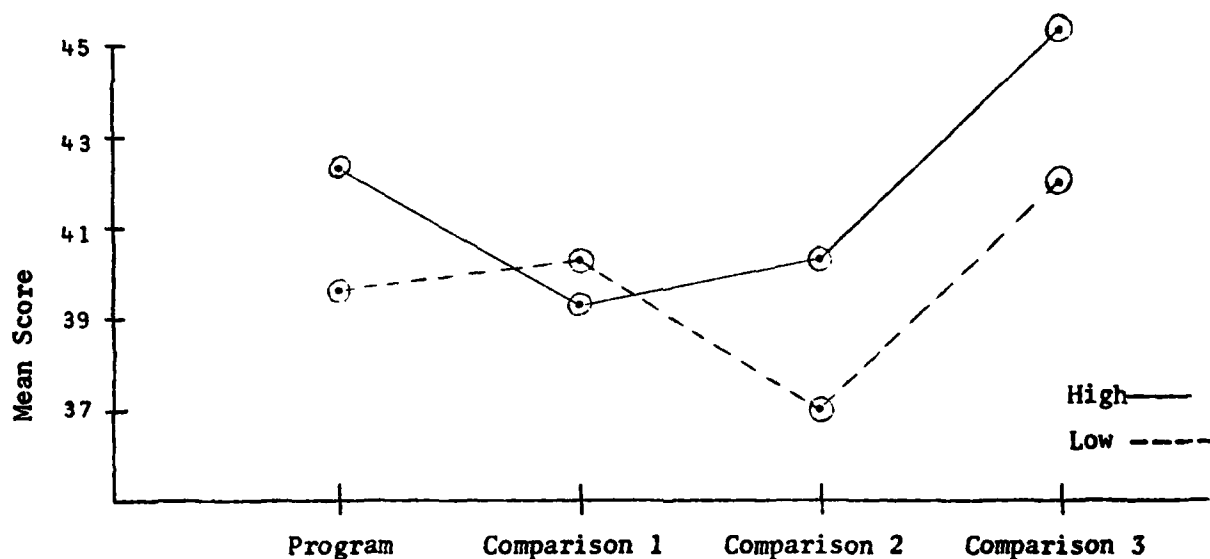


Figure 2. Cell means for the value of mathematics measure.

A test for interaction with the program school deleted yields an F-value of 1.51, which has a .23 level of significance. As such, the assumption of equality of the differences may be suspect, so the analysis outlined in section 3 is

applied. Application of equations (3.3) through (3.8) yield

$$\hat{\sigma}^2 = 25.8789, \quad \hat{\eta}^2 = .6192,$$

$$w_0^* = 8.9232, \quad w_1^* = 5.9736, \quad w_2^* = 5.3641, \quad w_3^* = 4.1033,$$

$$\delta' = 1.6182 \quad \text{and} \quad t^* = .7963.$$

The approximate level of significance of the statistic t^* with 2 degrees of freedom is $p = .2547$. That the level of significance is relatively high should be no surprise, for most of the significance can be attributed to the negative difference of comparison school 1.

5. CONCLUSION

In the analysis outlined in sections 2 and 3, deleting the program school then testing for interaction between school and ability at a high significance level is crucial. The entire analysis of section 2 hinges on the assumption that the average difference between high and low ability students in the absence of any program is constant across schools, and without that assumption, none of the results are justifiable.

The assumptions in section 3 that seem the most suspect are the assumption that the δ_i 's have a normal distribution, and the assumption that δ_0 has the same variance as the others. Clearly, some assumption on the distribution of the δ_i 's must be made, since m is typically small. That the distribution should be normal seems as reasonable as any, and makes the mathematics easy. The variance of δ_0 is unidentifiable, so the assumption of equality of variance seemed necessary.

Of course, even if the assumptions in section 2 or section 3 seem justifiable, the analysis above can never be considered as a replacement for a good experimental design with a control group. The assumptions necessary seem critical,

and minor deviations may cast doubt upon the conclusions. This analysis is suggested only when a control group is definitely unattainable.

Lemma 3.3 is clearly applicable to distributions on the δ_i 's other than the normal distribution, and as such an examination of other distributions may prove interesting. However, such alternate distributions usually are not employed in most designs employed by applied statisticians, so we leave this as an area of further research.

Of major interest is that under the same assumptions made in section 3, the statistic

$$\hat{t} = \frac{y_0 - m^{-1} \sum_{i=1}^m y_i}{[(m-1)^{-1} (\sum_{i=1}^m y_i^2 - m^{-1} (\sum_{i=1}^m y_i)^2)]^{1/2} [m^{-1}(m+1)]^{1/2}}$$

also has an asymptotic t distribution with $m-1$ degrees of freedom, and that in fact this can be more easily established by a straightforward appeal to the Skorohod convergence theorem (see Serfling (1980)). Therefore, the major justification of the use of the statistic in section 3 is the conjecture that it should outperform \hat{t} for small n_{ij} 's. This conjecture is an open question which could perhaps be answered by computer simulations or some other technique.

REFERENCES

Chung, K.L. (1974). A Course in Probability Theory, Academic Press, New York.

Graybill, F.A. (1976). Theory and Application of the Linear Model, Duxbury Press, Massachusetts.

Hersberger, J.R. (1983). The effects of a Problem Solving Oriented Mathematics Program on Gifted Fifth Grade Students, DAI 44A: 1715, Dec. 83, Ph.D. Thesis, Purdue University, W. Lafayette, IN.

Serfling, R.J. (1980). Approximation Theorems of Mathematical Statistics, J. Wiley and Sons, New York.

END

FILMED

3-86

DTIC